

·循证医学中的医学统计学问题·

## 如何规范地使用误差条形图

陈逸敏, 张晋昕

(中山大学公共卫生学院, 广州 510080)

**[摘要]** 误差条形图是一类应用广泛的统计图,既能用于统计描述,也能用于统计推断为目的的组间比较。实际应用中,有很多研究者错误地使用误差条形图进行统计推断。本文以两独立样本均值比较的例子,阐述误差条形图使用的误区,说明合理使用误差条形图的方式,旨在指导研究工作者规范地使用误差条形图。

**[关键词]** 误差条形图; 假设检验; 置信区间

**[中图分类号]** R195.1 **[文献标识码]** A **DOI:** 10.12019/j.issn.1671-5144.2017.05.015

## How to Use and Interpret Error Bars Properly

CHEN Yi-min, ZHANG Jin-xin

(School of Public Health, Sun Yat-sen University, Guangzhou 510080, China)

**Abstract:** Error bars commonly appear in figures in publications, which can give information about descriptive statistics, or inferential statistics. However, many researchers wrongly use error bars for statistical inference in practical applications. In this paper, we use examples of the comparison between two independent samples to illustrate the misuse of error bars and show the right way to use the error bars. Our aim is to guide the researchers to use error bars properly.

**Key words:** error bar; hypothesis test; confidence interval

误差条形图(error bar charts)是一种广泛应用于描述数据离散程度的统计图。误差条形图的主要使用形式有:①样本均值和个体的标准差(standard deviation, SD),②样本均值和标准误(standard error, SE),③置信区间(confidence interval, CI)。研究者常用误差条形图展示数据的变异或者不确定性,也将其辅助用于支持样本组间对比的统计推断结论。然而,许多研究者对误差条形图与统计推断间关系的理解存在误区<sup>[1-2]</sup>,甚至错误地完全依赖于误差条形图进行统计推断。本文将介绍误差条形图的用途,辨析误差条形图的误用。

### 1 误差条形图的主要用途

#### 1.1 统计描述

误差条形图可以提供统计描述的信息。在发表研究报告时,研究者不仅需要提供组间对比的定性结果,还应提供资料信息的定量描述,使结论更完整。常用 $\bar{x} \pm SD$ 的误差条形图呈现统计描述信息,直观地向读者展示各组的均值与变异程度,SD量化地呈现个体值的离散程度或者不确定性。

#### 1.2 统计推断

误差条形图还可以提供组间对比(假设检验)的信息。例如,置信度为95%的置信区间误差条形图中,误差线对应的区间 $[\bar{x}_1, \bar{x}_2]$ 之外是假设检验 $H_0: \mu = \mu_0$ 的拒绝域<sup>[3]</sup>。将某样本均值的置信区间与给定常数(如特殊总体的均值 $\mu_0$ )比较, $\mu_0$ 在95%置信区间的误差线以外,就等价于在0.05的显著性水平下拒绝零假设 $H_0: \mu = \mu_0$ ,从而判定组间差异有统计学意义<sup>[4]</sup>。

**[基金项目]** 中山大学教学改革研究重点资助项目(中大教务[2017]79-12-25)

**[作者简介]** 陈逸敏(1993-),女,浙江台州人,硕士研究生,从事统计学方法及其医学应用研究。

**[通讯作者]** 张晋昕, Tel: 020-87332453; E-mail: zhjinx@mail.sysu.edu.cn

## 2 误差条形图的误用

误差条形图可以辅助地反映统计推断的结论,但两者并不总是等价的。欲比较两组的均值是否相等时,误差条形图是否重叠并不等价于两组独立样本 $t$ 检验的 $P$ 值是否小于0.05。

研究者时常错误地解读误差条形图与组间对比(假设检验)之间的关系。国外一项研究纳入了473名作者在重要期刊发表的文章<sup>[2]</sup>,该研究表明31.5%的作者错误地认为两组间的均值误差条形图(包括均值标准误,95%置信区间)不重叠等价于两组的均值有显著性差异( $P<0.05$ )。可见,作者未加特别解释的误差条形图展示可能会让读者产生误解。

下文将用模拟研究和实例数据展示,两组之间对比时不宜用误差条形图取代两独立样本 $t$ 检验。

## 3 不宜用误差条形图取代假设检验的情形

模拟一:随机产生两组样本量各为500且服从正态分布的数据集。第1组的均值为50,标准差为10;第2组的均值为52.5,标准差为10。从第1组随机抽取100个数值作为样本I,从第2组随机抽取100个数值作为样本II。

欲比较样本I、II所来自总体的均值是否有差

异,采用两独立样本 $t$ 检验,结果为 $P<0.05$ (表1),提示两样本代表的总体均值存在显著性差异。从样本I、II的误差条形图(图1)中可以看出:①样本I、II的 $\bar{x}\pm SD$ 误差条形图的上边界与下边界重叠;②样本I、II的 $\bar{x}\pm SE$ 的误差条形图的上边界与下边界未重叠;③样本I、II的均值的95%置信区间有重叠。从模拟一可见,当两独立样本 $t$ 检验的 $P$ 值小于0.05时,误差条形图并不都是分离的。其中,两样本的 $\bar{x}\pm SD$ 误差条形图、均值的95%置信区间都重叠。说明这两类误差条形图分离不是两组间 $t$ 检验的 $P<0.05$ 的必要条件。所以,对这两类的误差条形图而言,“两组间的均值误差条形图未重叠等价于两组的均值有显著性差异( $P<0.05$ )”这一观点显然是错误的。该观点对于 $\bar{x}\pm SE$ 的误差条形图是否成立?本文用模拟二的例子来证明。

表1 样本I、II的基本信息

分组	样本量	均值	标准差	标准误	95%置信区间	两组间 $t$ 检验
I	100	49.82	10.26	1.03	(47.78, 51.86)	$P=0.02^*$
II	100	53.31	10.81	1.08	(51.16, 55.45)	

\* $P<0.05$ ,差异有统计学意义

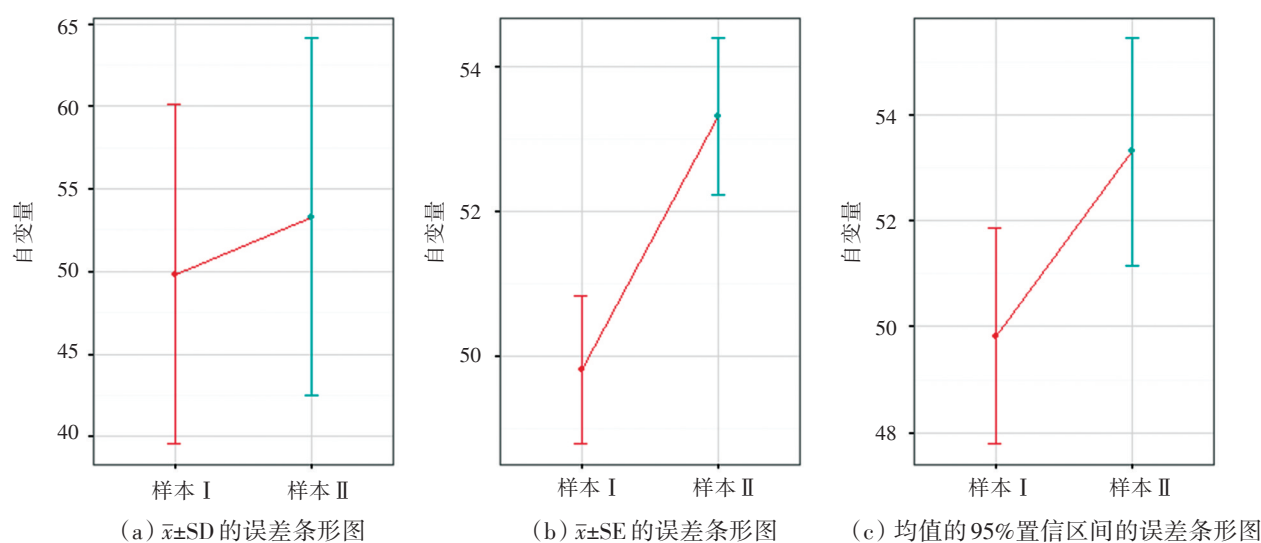


图1 样本I、II的三类常用误差条形图

模拟二:随机产生两组样本量各为500且服从正态分布的数据集。第1组均值为50,标准差为10;第2组均值为51.5,标准差为10。从第1组随机抽取100个数值作为样本III,从第2组随机抽取

100个数值作为样本IV。

欲比较样本III、IV所来自总体的均值是否有差异,采用两独立样本 $t$ 检验,结果为 $P>0.05$ (表2),提示两样本代表的总体均值不存在显著性差

异。但样本Ⅲ与样本Ⅳ的 $\bar{x} \pm SE$ 误差条形图的上下边界不重叠(图2)。据此,两样本的 $\bar{x} \pm SE$ 误差条形图分离并不是两组间 $t$ 检验得到 $P < 0.05$ 的充分条件。所以,对 $\bar{x} \pm SE$ 的误差条形图而言,“两组间的均值误差条形图未重叠等价于两组的均值有显著性差异( $P < 0.05$ )”这一观点也是错误的。

综合上述模拟一与模拟二的误差条形图可见:①模拟一、二中 $\bar{x} \pm SD$ 误差条形图的上边界与

表2 样本Ⅲ、Ⅳ的基本信息

分组	样本量	均值	标准差	标准误	95%置信区间	两组间 $t$ 检验
Ⅲ	100	49.81	10.26	1.03	(47.78, 51.86)	$P=0.10$
Ⅳ	100	52.31	10.81	1.08	(50.16, 54.45)	

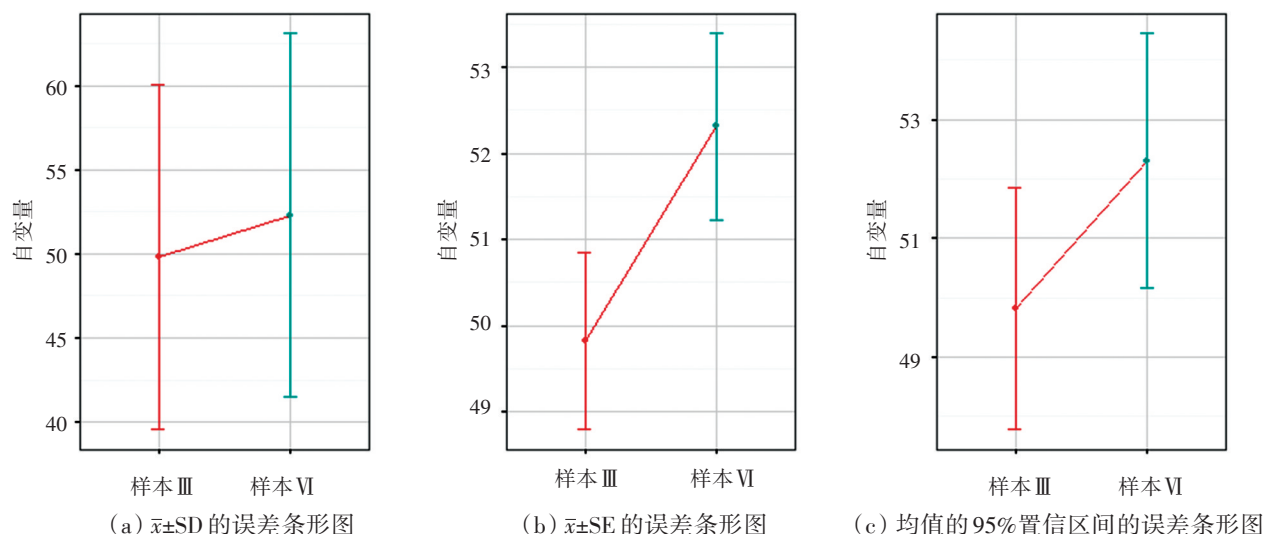


图2 样本Ⅲ、Ⅳ的三类常用误差条形图

下边界均重叠;②模拟一、二中 $\bar{x} \pm SE$ 误差条形图的上边界与下边界均未重叠;③模拟一、二中均值的95%置信区间均重叠。但模拟一的两组间 $t$ 检验 $P < 0.05$ ,而模拟二的两组间 $t$ 检验 $P > 0.05$ 。说明当两样本的某一类型误差条形图重叠(或者未重叠)时,对应假设检验的结果可能是 $P < 0.05$ ,也可能是 $P > 0.05$ 。所以,两组误差条形图是否重叠不宜直接取代两独立样本 $t$ 检验是否显著。

## 4 实例

在常用的误差条形图中,置信区间常被用于提供组间对比(假设检验)的信息。本文用下面的实例来展现95%置信区间是否重叠与两独立样本 $t$ 检验所得推论的关系。

某医生收集了25例正常人与25例冠心病患者的胆固醇(mmol/L)观察值,原始数据见表3<sup>[5]</sup>。

表3 观察25例正常人与25例冠心病患者的胆固醇水平

分组	胆固醇水平/(mmol/L)													
正常人	4.0	5.1	5.6	4.6	3.3	6.0	3.5	5.4	4.7	4.8	4.0	4.3	5.0	
	4.9	4.8	4.0	3.2	5.6	4.8	3.5	5.5	3.4	4.2	4.8	4.3		
冠心病患者	5.1	6.1	5.3	4.8	7.6	5.8	4.7	4.3	4.1	4.9	5.8	5.1	5.8	
	2.8	5.7	6.7	6.1	4.7	4.5	5.0	4.6	4.2	5.1	5.5	3.4		

现欲比较正常人与冠心病患者的胆固醇平均水平有无差别。

该数据可以用两独立样本 $t$ 检验进行组间比较,结果为 $t = -2.25$ ,  $P = 0.03 < 0.05$ ,说明正常人和冠心病患者的胆固醇均值差异有统计学意义。但从

图3可看出,正常人和冠心病患者胆固醇水平的95%置信区间有一定重叠。从该实例数据中可知两样本的95%置信区间重叠时,两组的均值也可能存在显著性差异(此例属于模拟一述及的情况)。

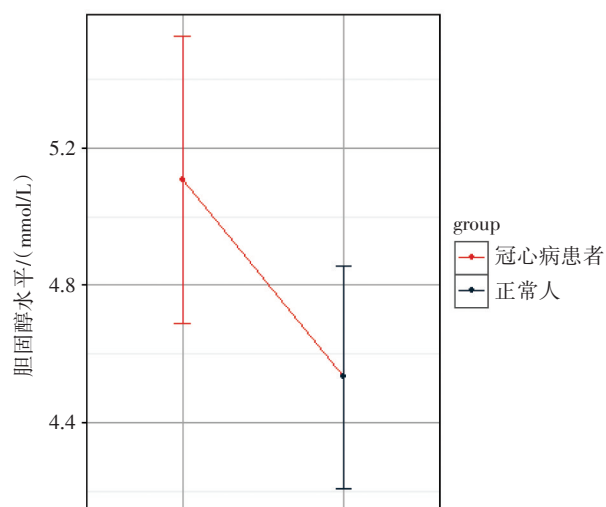


图3 正常人与冠心病患者的胆固醇水平及95%置信区间

## 5 讨论

为了强调整图通过误差条形图重叠与否便得出统计推断的对比结论是不妥的,本文着重论述了“误差条形图是否分离等价于假设检验的 $P$ 值是否小于显著性水平 $\alpha$ ”的矛盾之处。然而,误差条形图与假设检验的结果有时确是一致的。例如,本文提到两样本的 $\bar{x} \pm SD$ 误差条形图、95%置信区间分离都不是两组间 $t$ 检验的 $P < 0.05$ 的必要条件,需要注意的是这两类误差条形图分离是两组间 $t$ 检验的 $P < 0.05$ 的充分条件,所以这两类的误差条形图不重叠时,两组间 $t$ 检验的 $P < 0.05$ 。

为了避免出现错误的推论,请注意以下几点:

① 比较两样本均值是否有统计学差异,应以假设检验(如 $t$ 检验)的结果为准,辅以误差条形图呈现两样本直观的差异。② 基于两组数据加工出的误差条形图重叠(或者分离)一定程度的时候,确实会等价于组间对比的统计学结论。例如,两独立样本的样本量均大于10且为均衡设计时,两组均值的95%置信区间的误差条形图重叠小于1/4时,则有假设检验的 $P < 0.05$ <sup>[6-7]</sup>。在实际应用时,为了保证研究报告的可读性,不建议利用这种等价性试图用误差条形图代替统计学检验(或反之)。

## [参考文献]

- [1] CUMMING G, WILLIAMS J, FIDLER F. Replication, and researchers understanding of confidence intervals and standard error bars [J]. *Understanding Statistics*, 2004, 3 (4): 299-311.
- [2] BELIA S, FIDLER F, WILLIAMS J, et al. Researchers misunderstand confidence intervals and standard error bars [J]. *Psychol Methods*, 2005, 10(4):389-396.
- [3] 盛聚, 谢千式, 潘承毅. 概率论与数理统计[M]. 北京: 高等教育出版社, 2008:192-193.
- [4] 方积乾. 卫生统计学[M]. 北京: 人民卫生出版社, 2012: 116-118.
- [5] 方积乾. 生物医学研究的统计方法[M]. 北京: 高等教育出版社, 2007:422-423.
- [6] CUMMING G, FINCH S. Inference by eye: Confidence intervals and how to read pictures of data [J]. *Am Psychol*, 2005, 60(2):170-180.
- [7] CUMMING G, FIDLER F, VAUX D L. Error bars in experimental biology [J]. *J Cell Biol*, 2007, 177(1):7-11.

[收稿日期] 2017-06-01

(上接第296页)

- [16] MANDAL M, TRIPATHY R, PANDA A K, et al. Vitamin D levels in Indian systemic lupus erythematosus patients: Association with disease activity index and interferon alpha [J]. *Arthritis Res Ther*, 2014, 16(1): R49.
- [17] SMADJA D M, CORNET A, EMMERICH J, et al. Endothelial progenitor cells: Characterization, in vitro expansion, and prospects for autologous cell therapy [J]. *Cell Biol Toxicol*, 2007, 23(4):223-239.
- [18] WERNER N, NICKENIG G. Clinical and therapeutical implications of EPC biology in atherosclerosis [J]. *J Cell Mol Med*, 2006, 10(2):318-332.
- [19] LEE P Y, LI Y, RICHARDS H B, et al. Type I interferon as a novel risk factor for endothelial progenitor cell depletion and endothelial dysfunction in systemic lupus erythematosus [J]. *Arthritis Rheum*, 2007, 56(11): 3759.
- [20] DENNY M F, THACKER S, MEHTA H, et al. Interferon -

alpha promotes abnormal vasculogenesis in lupus: A potential pathway for premature atherosclerosis [J]. *Blood*, 2007, 110 (8): 2907-2915.

- [21] WESTERWEEL P E, LUIJTEN R K, HOEFER I E, et al. Haematopoietic and endothelial progenitor cells are deficient in quiescent systemic lupus erythematosus [J]. *Ann Rheum Dis*, 2007, 66(7): 865-870.
- [22] ADLER B, ADLER H, JUNG T W, et al. Interferon -alpha primes macrophages for lipopolysaccharide-induced apoptosis [J]. *Biochem Biophys Res Commun*, 1995, 215(3): 921-927.
- [23] BOTELLA - CARRETERO J I, ALVAREZ - BLASCO F, VILLAFRUELA J J, et al. Vitamin D deficiency is associated with the metabolic syndrome in morbid obesity [J]. *Clin Nutr*, 2007, 26(5):573-580.

[收稿日期] 2017-04-04